Predicting the Probability of *Tough* Construction from the *Tough* Object's Syntactic Category

Myoung Hyoun Song
(Seoul National University)

Song, Myoung Hyoun. 2009. Predicting the Probability of *Tough* Construction from the *Tough* Object's Syntactic Category. *SNU Working Papers in English Linguistics and Language 8, 28-42.* This paper sets forth the syntactic alternation between the *tough* construction (TC, e.g. The bear is easy to see.') and the *tough*-related extraposition construction (EC, e.g. It is easy to see the bear.'). The TC-EC alternation shows proportional difference in the overall ratio of TC to EC (4:6). This difference is accounted for by syntactic factors such as 'Syntactic Category', 'Modification'. The odds ratio of TC to EC in the 'Syntactic Category' was 8.9:1.1 for 'Pronoun', 4.2:5.8 for 'DP', 1.4:8.6 for 'NP', and 0:10 for 'CP'. The odds in the 'Modification' was 6.4 to 3.6 for 'Head-only', 3.3:6.7 for 'Premodified', 0:10 for 'Postmodified'. With the 'Syntactic Category' remaining as the only single predictor variable of the probability of the TC, a pronoun *tough* object is 45.5 times more likely to occur in TC than an NP *tough* object, 10.5 times than a DP *tough* object. These findings were obtained by a chi-square test of independence and a logistic regression analysis with 100 *tough*-related utterances from the spoken texts of on-line corpus COCA. **(Seoul National University)**

Keywords: *Tough* Construction, Extraposition Construction, Syntactic Alternation, Chi-square, logistic regression

1. Introduction

The *tough* construction (hereafter TC; *tough* movement (Chomsky 1976), Null Operator Construction (Browning 1991), Unbounded Dependency Construction (Bouma 2001)) is defined as a construction in which a *tough* adjective[1] takes a subordinate infinitive clause missing an object, as

1) *Tough* adjectives are classified into five distinct categories in terms of their semantic contents as in the followings (Song 2004).
   a. Difficulty : awkward, difficult, easy, hard, impossible, simple, tough, etc.
   b. Danger : dangerous, hazardous, risky, safe, unsafe, etc.
   c. Value : awful, convenient, excellent, fine, good, harmful, horrible, nice, etc.

illustrated in (1).[2)]

(1) a. For some people, the bear is easy to see.
b. Driver's licenses remain very easy to get with little, if any, checking.
c. They're very easy to transport.
d. What Gary is doing would be easy to miss.

What the underlined phrases in (1) have in common is that the *tough* adjective 'easy' takes a subordinate infinitive clause of the syntactic composition [CP Ø-OPi C [IP I [VP V ti]]] (Browning 1991). The trace left in the missing object (or object gap in (Bouma et al. 2001)) is anyhow connected to the matrix subject for the interpretation of the whole construction (Dalrymple 2001). The connectivity between the matrix object and the missing object (or *tough* object) is explicitly witnessed in the instances where the TC alternates with the *tough*-related Extraposition Construction (hereafter, EC; pleonastic/sentential subject construction (Kim 1996), Anticipatory-it Construction (Mair 1987)), as illustrated in (2).

(2) a. For some people, it is easy to see the bear.
b. It remains very easy to get driver's licenses with little, if any, checking.
c. It is very easy to transport them.
d. It would be easy to miss what Gary is doing.

The EC sentences in (2) are by-products of syntactic alternation from the corresponding TC sentences in (1). Specifically, the *tough* object of definite NP (or DP) 'the bear' in (2a) can occur freely either in EC or TC. The *tough* object of bare plural NP 'driver's licenses' in (2b) can also be realized either in EC or TC. The same structural freedom is true of the *tough* object of pronoun 'them' in (2c) and even the clause in (2d).

---

d. Interest : delightful, fun, pleasant, etc.
e. Stimulation : amazing, annoying, boring, exciting, stimulating, etc.

2) The illustrative examples in this paper are direct search results (or sometimes adapted) from the Corpus of Contemporary American English, the 365 million word corpus, which is available at http://corpus.byu.edu.

According to Song (2004), the odds ratio of TC is 4:6, compared to the EC, especially when a *tough* object occurs after the *tough* adjective 'easy'. The ensuing question is "What factors cause the difference in odds ratio between the TC and the EC? More specifically, what factors contribute to the occurrence of the TC and what factors to that of the EC?

Mair (1987) addressed that question, quoted as saying "The corpus contains numerous examples in which *tough*-Movement creates a situationally appropriate distribution of given and new information in a sentence. (p. 63)" He seemed to propose a single factor 'Givenness' in the pragmatic or information structure dimension as a main effect on the odds of TC and EC. This means that when a speaker refers to information given in the previous discourse or assumed to be in the common ground, s/he packs that information into the matrix subject in the TC with a pronominal form ('it', 'they', etc.) or a deictic form ('this', 'that', 'these', those', etc.). But when a speaker feels the need to introduce new information into the common ground, s/he uses the EC by packing that information with a focus-marked object usually at the end of the utterance. With this 'Givenness' factor, it would be incorrectly predicted that only the DPs and the pronouns in (1) can occur in the TC, but the bare plural NPs or the clauses in (1) can't, since the latter ones usually serve to introduce new information. Moreover, it is wrongly predicted that the *only*-modified object can't occur in the matrix subject position of TC, since it typically introduces a focused semantic value with a set of alternatives, together with a normal semantic value, into the common ground (Rooth 1992), which is not borne out with the underlined example in (3)

(3) For starters, the election year math is decidedly in their favor; of the 34 seats up for re-election, 22 are held by Republicans. Of the remaining dozen Democratic-held seats, only a couple will be difficult to defend: Sens. Mary Landrieu of Louisiana and Tim Johnson of South Dakota could face stiff challenges.

The *only*-modified matrix subject 'only a couple' in (3) occurs in the TC, even though it typically denotes new information not given in the previous discourse. Bresnan (2007) proposed an assortment of factors from many linguistic dimensions, that is, pronominality from syntax,

definiteness and animacy from semantics, givenness from pragmatics, argument length difference from quantitative statistics to investigate into the odds of dative alternation. Her conclusion was that "nongiven or indefinite themes and pronominal recipients favor V NP NP, pronominal themes and indefinite recipients favor V NP PP. She tried to explain the syntactic phenomenon of dative alternation with as many factors as the linguistic knowledge permits, but she overlooked the side-effects of a composite of diverse factors, that is, multi-collinearity effects. We can see at first glance that the predictor variables are highly correlated between each other. A pronoun (from Pronominality), for example, is highly correlated with a definite NP (from Definiteness) in the sense that they both refers to a specific individual or sometimes property. And the definite NP does, in turn, show a high correlation with a given expression (from Givenness) in the sense that they commonly convey information already given in the common ground. The same correlation is true for other predictor variables. The multi-collinearity effects between predictor variables in Bresnan (2007) were certain to distort the significance of the effects. She claimed, for instance, that the main effects of 'non-given recipient' was not statistically significant, but she could have got a significant result with uncorrelated variables in her experiment.

The purpose of this paper is to investigate the relationship between un-correlated predictors only from the syntactic dimension and the odds ratio of the TC-EC alternation. Given that the TC-EC alternation is basically a syntactic phenomenon, it makes sense to explain that alternation only with syntactic factors. As above, Mair (1987)'s pragmatic approach to the syntactic phenomenon was limited in encompassing all the TC instances. And Bresnan (2007)'s mixed-effects approach was problematic with high correlation between the predictors. The alternative option is to approach the syntactic phenomenon with only syntactic factors based on the unrelated features, such as 'syntactic category', 'plurality', and 'modification' (See details in 2.2). And this paper is also designed to predict the probability of the occurrence of TC from the selected predictors, compared to the occurrence of EC. The hypotheses for this study are stated as follows:

1. There is a relationship between syntactic features of the *tough* object and the odds of TC to EC.
2. The probability of the TC is predicted from the syntactic features,

especially the syntactic category of the *tough* object.

2. Method

The data for analysis were obtained from search results of the online Corpus of Contemporary American English (COCA, available at http://www.americancorpus.org/) and put into chi-square tests and logistic regression analysis to support the research hypotheses in the introduction. We selected the online COCA corpus, since it is one of the largest database of written and spoken texts, totaling 380 million words, and it is freely available online.

2.1 Data collection

As mentioned in the introduction, the overall odds ratio of TC:EC was 4:6. But the ratio varies among *tough* adjectives. One *tough* adjective, 'dangerous' showed a bias towards the EC with the odds ratio of 1:9 out of the total number of 83 tokens while another *tough* adjective, 'simple' had a bias towards the TC with 7:3 out of 64 tokens (Song 2004). This variance had us select the *tough* adjective 'easy', which showed almost the same ratio as the overall ratio. And we limited the scope of search to the data from spoken texts, since those data are sufficient to see a general tendency of the TC-EC alternation. And we limited the data only to the utterances in which the *tough* adjective 'easy' immediately precedes the subordinate infinitive clause. The relevant search results amounted to 2,659 utterances in total and they were sorted by frequency from recent years to past years. Then we designed a randomly selected dataset by picking out the first 10 relevant utterances from each year over a ten-year period from 2008 to 1999. We mean the relevant utterances by the data containing the *tough* adjective 'easy' which is available for TC-EC alternation with a transitive verb in it. In addition, we intended a balanced dataset of 2 utterances from each of the broadcast media, including ABC GMA, CBS 48 Hours, NPR Talk Net, Fox HC, and NBC Today. After all, we obtained a dataset of *tough*-related utterances, as shown in Table 1.

Table 1. A dataset of *tough*-related utterances from the spoken text of COCA

Years	Text Type	Broadcast Media	No. of Utterances	Total No. of Utterances
2008 - 1999	Spoken	ABC GMA	20 (2/year)	100
		CBS 48 Hours	20 (2/year)	
		NPR Talk Net	20 (2/year)	
		CNN Intl	20 (2/year)	
		NBC Today	20 (2/year)	

2.2 Data analysis

Data analysis and a coding procedure were conducted in two stages. On the first stage of analysis, we decided on whether each utterance belongs to TC or EC. The EC utterance was coded with 0 and the TC utterance with 1. The coding method was good for converting the EC/TC variable into a dichotomous variable, which was submitted into subsequent statistic tests as a dependent variable. On the second stage, we looked through the *tough* object (the matrix subject in TC and the subordinate object in EC) in each utterance for the distinctive syntactic features such as a syntactic category, plurality, pre/post-modification. As for a syntactic category, the *tough* objects are classified into 4 subcategories, which are illustrated in (4).

(4) a. DP: definite and specific NP (e.g. the bear, his hotel, Patty's killer, these pieces)
b. NP: indefinite and nonspecific NP (e.g. an American president, driver's licenses, somebody who's vulnerable)
c. Pronoun: referential anaphora (e.g. it, they, we, him, that)
d. CP: that-clause, WH-clause (e.g. that Adams were getting, why drive-ins are making···)

We accepted the distinction between a DP and an NP in Bowers (1988) since we observed that a nominal phrase differs in their preference for the TC with respect to the definiteness /specificity of its specifier. And

we observed that the *tough* object of the subcategory CP opted for the EC without exceptions. Thus we ruled out the subcategory CP from our analysis and coded the rest into three ways : NP=1; DP=2; Pronoun=3. This tripartite division allows the syntactic category to be a categorical variable, which is subsequently submitted to a logistic analysis as a predictor. And as for 'plurality', the *tough* objects are classified into 'Singular' subcategory and 'Plural' subcategory, which are coded with 0 and 1 respectively as another dichotomous predictor. The distinction with respect to 'plurality' was due to the observation that some singular NPs (e.g. a building, an American president) opted for the EC while other plural NPs (e.g. driver's licenses, endangered species) showed preference for the TC. Lastly, as for 'modification', the *tough* objects are classified into three distinct categories such as 'Head only', 'Pre-modified' and 'Post-modified', as illustrated in (5).

(5) a. Head only : Head N (e.g. a building, the country, despair)
b. Pre-modified: modifier + Head N (e.g. fake documents, flat abs, your common sense, adult bedbugs)
c. Post-modified : Head N + modifier (e.g. a permit to carry a concealed…, the facts on that one, those situation you just saw.)

The three-way distinction (Head-Only (1), Pre-modified (2); Post-modified (3)) was later reduced to the two-way distinction by incorporating the 'Pre-modified' category and the 'Post-modified' category into a combined category since the preference for the TC was contingent on the existence of any modifiers. Thus, the 'Head-only' category was coded with 0 and the 'Modified' category into 1, which creates one more dichotomous predictor. The coding system is illustrated in (6).

(6) a. police and her parents were confident Patty's killer would be easy to find. (2008, CBS-48 Hours)
→ Coded as : TC (1); DP (2); Singular (0); Head Only (0)
b. It's relatively easy to get to places like Fuket, in Thailand. (2005, CNN-Intl)
→ Coded as : EC (0); NP (1); Plural (1); Modified (1)

The coded data ware imported into the SPSS (17th version) for analysis. First the chi-square test was used to investigate the independence of the predictor variables from the dependent variable at the alpha level of significance 0.05. Then the logistic regression analysis was used to predict the probability of the TC from the predictor variables. Unlike two other predictors, the 'Syntactic Category' is a categorical variable with three levels. Thus we set up two dummy variables with the 'Pronoun' subcategory as a reference group for comparison.

3. Results

We've got results from the *tough*-related data analysis to partly support the hypotheses in the introduction, repeated below.

1. There is a relationship between syntactic features of the *tough* object and the odds of TC to EC.
2. The probability of the TC is predicted from the syntactic features, especially the syntactic category of the *tough* object.

3.1 Independence of syntactic variables from the odds of TC

The chi-square tests of independence revealed that the 'Syntactic Category' variable and the 'Modification' variable are significantly associated with the odds of the TC but the 'Plurality' is not. Specifically, the relationship between the 'Syntactic Category' and the dependent variable is shown in the Table 2.

Table 2. Relationship between Syntactic Category and EC/TC

		CP	NP	DP	PRO	Total
EC	Count	11	24	15	4	54
	Percentage	100.0%	85.7%	57.7%	11.4%	54.0%
TC	Count	0	4	11	31	46
	Percentage	.0%	14.3%	42.3%	88.6%	46.0%
Total	Count	11	28	26	35	100
	Percentage	100.0%	100.0%	100.0%	100.0%	100.0%

The x2 statistic (46.387, df=3, p=.000) suggests that there is a significant association between the 'Syntactic Category' variable and the Odds ratio of TC. The overall ratio of TC to EC is 4.6 : 5.4, which suggests that the syntactic category of *tough* objects shows a preference for EC. This general tendency is strengthened by CP *tough* objects. They show the odds ratio of TC to EC is 0 to 100. This means that a that-clause or WH-clause *tough* object is 100% likely to occur in EC, as in (7a). NP *tough* objects show the odds ratio of 1.4:8.6, which means that a majority of indefinite NPs are likely to occur in the EC, as in (7bi). Note that the exceptional cases that occur in TC are modified bare plurals, as in (7bii), or idiomatic expressions, as in (7biii). And DP *tough* objects show the odds ratio of 4.2:5.8, which suggests that definiteness or specificity of NP doesn't cause a big difference in distribution. Rather, the DPs with modifiers show preference for the EC while the DPs without modifiers opt for the TC, as in (7ci) and in (7cii) respectively. Lastly, pronoun *tough* objects show the odds ratio of 8.9 : 1.1, as in (7dii), which suggests that almost all pronouns are likely to occur in TC with some exceptional cases in which the *tough* adjective 'easy' doesn't refer to the degree of difficulty, but it refers to psychological state as in (7di). Interestingly enough, an only-modified *tough* object occurs in EC, as in (7e), which agrees with Mair (1987)'s claims on 'information packaging'.

(7) a. It's very easy to deduce that unless every single terrorist in the world, the tens of thousands ···

b. (i) It's easy to shut down a building, but as you see in this graphic, it ···

(ii) Endangered species are just as easy to hit as a would-be land mine.

(iii) Well, it's easy to point fingers.

c. (i) It will be easy to rid the panel of jurors who are blatantly political in their views of things ···

(ii) 95% of people said that the machines were easy to use.

d. (i) And we are trying our level best. It was not easy to lose him.

(ii) They are very easy to transport.

e. It's easy to see only ruin.

The relationship between the 'Modification' and the dependent variable is shown in the Table 3. Only *tough* objects of the syntactic category DP, NP and Pronoun are included in the 'modification counts' (89 utterances in total), minus the CP objects (11 utterances) which aren't usually modified.

Table 3. Relationship between Modification and EC/TC

		Head Only	Pre-modified	Post-modified	Total
EC	Count	24	6	13	43
	Percentage	35.8%	66.7%	100.0%	48.3%
TC	Count	43	3	0	46
	Percentage	64.2%	33.3%	.0%	51.7%
Tota l	Count	67	9	13	89
	Percentage	100%	100%	100%	100%

The x2 statistic (19.309, df=2, p=.000) suggests a significant association between the 'Modification' variable and the Odds ratio of TC. The overall ratio of TC to EC is 5.2 : 4.8, which suggests that the modification of *tough* objects shows a preference for TC. This general tendency is caused by the overwhelming frequency of head-only *tough* objects over modified *tough* objects. First, the head-only *tough* objects opt for the TC with the odds ratio of TC to EC (6.4:3.6), as in (8ai). Exceptionally, an abstract noun as in (8aii) or a deixical pronoun 'that' as in (8aiii) occurs in the EC. And a majority of pre-modified *tough* objects favor the EC with the ratio of 3.3 : 6.7, as in (8bi) and (8bii) respectively while all of the utterances involving post-modified *tough* objects occur in EC, as in (8c).

(8) a. (i) The bear is easy to see.
.(ii) Well, because it's easy to write about style.
.(iii) It's not easy to do that simply from the air.
b. (i) Cooper says adult bedbugs are easy to spot.
.(ii) It's very easy to get flat abs.
c. It's relatively easy to get a permit to carry a concealed weapon,

And the relation between 'Plurality' variable and the odds ratio of

TC/EC is nonsignificant (x2 (d.f.=1, N=89) = .000, p > .05). This means that it doesn't matter whether a *tough* object is singular or plural in explaining the difference in odds ratio between TC and EC.

Summing up, two of the three syntactic variables (Syntactic Category and Modification) are significantly associated with the odds ratio of EC to TC whereas the other variable (Plurality) is not associated with the dependent variable. A majority of NP *tough* objects tend to prefer the EC, while the Pronoun *tough* objects tend to prefer the TC. DP *tough* objects are slightly biased towards the EC. And the head-only *tough* objects prefer the TC whereas the modified *tough* objects prefer the EC. The variable 'Plurality' has no significant effect on the odds ratio of TC to EC. With these significant variables, we expected to predict the probability of the TC with logistic regression analysis, which is dealt with in the following subsection.

3.2 Predicted probability of the TC

Logistic regression analysis was employed to predict the probability that a *tough* object would occur in TC from predictor variables. The predictor variables were 'Modification', and 'Syntactic Category'. The number of relevant cases is 89, with an exception of the CP *tough* objects, which are not modified. We selected the Forward method to enter the predictor variables, since we wanted to find out the strongest predictor with statistical significance. The Syntactic Category variable was dummy-coded, with the Pronoun as the reference group. Table 4 shows the results of the logistic regression analysis with the variable entered in the first step.

Table 4. Logistic Regression Predicting the Probability of TC from Syntactic Category.

	B	S.E.	Wald	df	Sig.	Exp(B)
Syntactic Category			26.418	2	.000	
NP	3.839	.758	25.685	1	.000	.022
DP	2.358	.663	12.640	1	.000	.095
Constant	2.048	.531	14.855	1	.000	7.750

The Syntactic Category was entered in the first step, improving the model with x2= 40.010, p = .00. (d.f.=2, N=89) The model was able to correctly classify 67.4% of the *tough* objects in TCs and 90.7% of the of the *tough* objects in ECs with an overall success rate of 78.7%. Employing the alpha significant level of 0.05, the NP *tough* objects and the DP *tough* objects significantly occur in TC less than the Pronoun *tough* objects, the reference group. The exponentiated coefficient (Exp(B)) 0.022 of the NP *tough* objects indicates that they are 0.022 times more likely to occur in TC than the Pronoun *tough* objects. The inverted odds ratio '45.5' suggests that Pronoun *tough* objects are 45.5 times more likely to occur in TC than the NP *tough* objects. In the same way, the DP *tough* objects have an odds ratio of .095 (inverted to 10.5), which suggests that Pronoun *tough* objects are 10.5 times more likely to occur in TC than the NP *tough* objects. Suppose there are three phrases such as 'the ball', 'a ball', 'it' which refers to the same entity in the universe of discourse. If they occur in a *tough*-related construction, the odds ratio of 'it' occurring in TC (e.g. It is easy to stop.) is 10.5 times higher than the DP 'the ball' (e.g. The ball is easy to stop.), 45.5 times higher than the NP 'a ball' (e.g. A ball is easy to stop.), . The 'Modification' predictor failed to enter the model (d.f.=1, p>0.05)

4. Discussion

The motivation to start this research was that we should explain a syntactic phenomenon such as 'TC-EC alternation' with syntactic factors, rather than other factors in different dimensions of linguistics. The results were favorable to our efforts. The overall odds ratio of TC to EC (4.6 : 5.4) in this study is similar to the ratio (4:6) of the previous study (Song 2004). It's noteworthy that even a small sample of 100 *tough*-related utterances can show a similar alternation pattern to a lot larger sample of almost 1,000 sentences in the previous study. This means that the syntactic alternation between the TC and the EC is regular, not incidental, among the English speakers. This regular syntactic alternation was accounted for by syntactic features such as 'Syntactic Category', 'Modification', and Plurality which all involve the *tough* object. To begin with, the variable 'Syntactic Category' and the variable 'Modification' were significantly associated with the odds of TC. The odds ratio of TC

to EC in the 'Syntactic Category' was 8.9:1.1 for 'Pronoun', 4.2:5.8 for 'DP', 1.4:8.6 for 'NP', and 0:10 for 'CP'. Given that the pronoun and the DP has a semantic type of *e*, denoting an individual while the NP and the CP has a semantic type of <*e,t*> and <*st,t*> respectively, denoting a property and a truth value, we can generally say that the *e* type *tough* objects opt for TC whereas the <*e,t*> or <*st,t*> type *tough* objects prefer EC. And the variable 'Modification' were also significantly associated with the odds of TC. The odds in the 'Modification' was 6.4 to 3.6 for 'Head-only', 3.3:6.7 for 'Pre-modified', 0:10 for 'Post-modified'. This suggests that the existence of the modifiers in *tough* objects tip the balance of the odds ratio of the TC-EC.

With the 'Syntactic Category' remaining as the only single predictor variable of the probability of the TC, a pronoun *tough* object are 45.5 times more likely to occur in TC than an NP *tough* object, 10.5 times than a DP *tough* object. One contribution of this paper is that we propose 'Syntactic Category' as a predictor variable to choose between the EC and the TC. As mentioned in the introduction, the 'Givenness' factor in Mair (1987) is limited in the sense that it encompasses only a DP *tough* object and a pronoun *tough* object in the matrix subject position of the TC. A composite of factors in Browning (2007) has an implicit problem with multi-collinearity effects. The proposed 'Syntactic Category' factor encompasses all the possible *tough* objects no matter whether it occurs in TC or EC and causes no multi-collinearity problems with any other syntactic variables.

This study is limited in three ways. First, it limited its scope only to the spoken data involving the 'easy' *tough* adjectives. It would be beneficial to conduct a comparative study of the change in the odds ratio of TC-EC in terms of the text types (written vs. spoken). And the generalization problem would be overcome with a comprehensive study dealing with all the *tough* adjectives, which was contained in Song (2004). Secondly, the syntactic features were limited to the three variables. Actually, the argument structure in which a *tough* object occurs can be a deciding factor. It means that a *tough* object with one-place transitive verb is more likely to occur in TC than with two-place transitive verb or with the following modifier. This variable was thrown away, since the dataset for this research doesn't give enough information on argument structure involving the *tough* object. With more time and space, we could design an adequate (hopefully elegant) research with more syntactic

variables. Finally, the intended syntactic account of the syntactic phenomenon itself is limited. Even though the TC-EC alternation is assumed to be a syntactic phenomenon, it doesn't preclude any other approaches in the linguistic dimensions. That is, it is accessible semantically, pragmatically, or informationally. It would be good to see interaction effects between the predictor variables in each dimension of linguistics, as long as we avoid the multi-collinearity effects.

5. Conclusion

The TC-EC alternation shows difference in odds ratio in spoken texts of broadcast media. That proportional difference can be significantly accounted for by syntactic factors such as 'Syntactic Category' and 'Modification. Above all, the 'Syntactic Category' was the only single predictor to estimate the probability of the TC. The pronoun is stronger than the DP, which is stronger than the NP in predicting the probability of the TC. This syntactic account of the proportional difference in the TC-EC alternation will serve as a guide to various linguistic research to other structural alternations.

References

Bouma, et al. 2001. Satisfying Constraints on Extraction and Adjunction. *NLLT 19:* 1-65.

Bowers, J. 1988. Extended X-Bar Theory, the ECP, and the Left Branch Condition. *The Proceedings of the West Coast Conference on Formal Linguistics (WCCFL) 7*: 47-62. Stanford Linguistics Association, Stanford University.

Bresnan, J. 2007. Is syntactic knowledged probablistic? Experiments with the English dative alternation. CSLI. Stanford.

Browning, M. A. 1991. Null Operator Constructions. *Outstanding Dissertations in Linguistics.* In J. Hanamer(ed.), New York: Garland Publishing, inc.

Chomsky, N. 1976. On Wh-Movement. Mimeo. Printed version. In P. W. Culicover, et al.(eds.), *Formal Syntax,* 1977: 71-132. New York: Academic Press.

Dalrymple and King. 2000. Missing-Object Constructions: Lexical and Constructional Variation. *Proceedings of the LFG Conference*: 1-22. In M. Butt and T. H. King(eds.), CSLI Publications.

Kim, Boomi. 1996. Predication in *Tough*-Construction. *The Proceedings of the Fourteenth*

West Coast Conference on Formal Linguistics: 271-285. In J. Camacho, L. Choueiri and Maki Watanabe (eds.), CSLI Publications.

Mair, C. 1987. *Tough*-movement in present-day British English: A corpus-based study. *Studia Linguistica 41 (1):* 59-71.

Rooth, M. 1992. A theory of focus interpretation. *Natural Language Semantics*: 75-116.

Song, M. 2004. *A Corpus-based Analysis of the English Tough Construction.* M.A. thesis, Seoul National University. Published by Verlag Doktor Mueller (VDM) in 2008.

Myoung Hyoun Song
y720four@yahoo.com